consensus sequence. The mismatching nucleotides, however, have a more destabilizing effect than the ones found in the first sequence. The S2 subunit ORF is not closely preceded by a ribosomal binding sequence, which may suggest that S2 is translated through a mechanism not involving the detachment and reattachment of the ribosome between the coding regions for S1 and S2. The short distance between the S1 and S2 cistrons, and the absence of a ribosomal binding site are characteristic of this mechanism. A ribosomal binding site for S4 in the sequence CAGGGCGGC (SEQ ID NO:24), starting at position 2066 is possible. The ORF for S5 is preceded by the sequence AAGGCG (SEQ ID NO:25), starting at position 2485, which matches five out of six nucleotides in the consensus sequence AAGGAG (SEQ ID NO:26). Finally, S3 is preceded by the sequence GGGAACAC (SEQ ID NO:27), which is very similar to the proposed ribosomal binding site for S1, i.e., GGGAAGAC (SEQ ID NO:28).--

REMARKS

The Notice to Comply with Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Sequence Disclosures dated July 2, 2002 indicates that the application failed to comply with the requirements of 37 C.F.R. §§1.821-1.825. Accordingly, applicants are required to submit an initial Sequence Listing in computer readable form, and a statement that the computer readable and paper forms of the Sequence Listing are identical.

As a result of this Amendment, the paragraphs at page 9, lines 20-26 of Table 1; at page 14, line 35 through page 15, line 12; the Table, on page 21; and the paragraphs, at page 26, line 22 through page 27, line 25; and at page 28, line 15 through page 29, line 12 were amended to include the appropriate SEQ ID NOS. This initial Sequence Listing is submitted, herewith, in both paper and computer readable forms, and is accompanied by a statement indicating that the paper and computer readable forms are identical.

Finally, as required by 37 C.F.R. §1.121, a marked up version of the replacement

paragraphs and table of the specification is attached with additions indicated by underlining and deletions indicated by brackets.

It is believed that, as presently amended, the instant application fully complies with the requirements of 37 C.F.R. §§1.821-1.825.

AUTHORIZATION

No fee is believed due.

Should any fee(s) be required by the filing of this Amendment and accompanying papers, authorization is hereby given to the Commissioner to charge the amount of any such fee(s) that is/are properly assessable in this application to Deposit Account No. 13-4500, Docket No. 2026-4253US7. A DUPLICATE COPY OF THIS SHEET IS ATTACHED.

Respectfully submitted,
MORGAN & FINNEGAN, L.L.P.

Dated: July 24, 2002

By: Dorothy R. Auth

Dorothy R. Auth
Registration No. 36,434

Correspondence Address:
MORGAN & FINNEGAN, L.L.P.
345 Park Avenue
New York, NY 10154-0053
(212) 758-4800 Telephone
(212) 751-6849 Facsimile

VERSION WITH MARKINGS TO SHOW CHANGES MADE

As required by 37 C.F.R. §1.121, a marked up version of the replacement claims is as follows:

Please amend and replace the paragraph, at page 9, lines 20-26 of Table 1, as follows:

--The S4 H_2N-terminal amino acid sequence determined using the automated protein sequenator is shown in blocks as the mature protein sequence. Residues that were questionable in the sequence are indicated by brackets. The DNA (SEQ ID NO:1) and predicted amino acid sequences (SEQ ID NO:2) are shown. Possible initiation codons are indicated by f-Met. A putative proteolytic cleavage site is indicated by *. The mature protein sequence as indicated is SEQ ID NO:4. The oligonucleotide probe sequence is shown in the block labeled probe 21D3 (SEQ ID NO:3). The abbreviations used are: P = G or A; Y = T or C; N [-] = A, C, G, or L.--

Please amend and replace the paragraph, at page 14, line 35 through page 15, line 12, as follows:

--Because of the high C+ G content of B. pertussis DNA, it was necessary to use both of the above mentioned methods with a combination of 8% and 20% polyacrylamide- 8 M urea gels for sequence analysis. Each nucleotide has been sequenced in both directions on average of 4.13 times. The final consensus sequence of the sense strand is shown in Table 2. It is noted that the sequence of the S4 subunit gene has been included in this table for completeness since this sequence lies in the middle of the structural gene sequence presented in Table 2. The complete nucleotide sequence of Pertussis Toxin gene (SEQ ID NO:5) and deduced amino acid sequence (SEQ ID NOs:6-11) are presented in Table 2. The entire sequence contains about 62.2% C+G with about 19.6% A, 33.8% C, 28.4% G and 18.2% T in the sense strand, wherein A, T, C and G represent the nucleotides adenine, thymine, cytosine and guanine, respectively.--

Please amend and replace the Table, on page 21, as follows:

--

Table 4

Comparison of Two Homologous Regions in ADP-ribosylating subunits of Pertussis, Cholera, and E. Coli Heat Labile Toxins

Region 1	SEQ ID NO:										
Pertussis S1 subunit	12	(8)	Tyr	Arg	Tyr	Asp	Ser	Arg	Pro	Pro	(15)
Cholera[4] A subunit	13	(6)	Tyr	Arg	Ala	Asp	Ser	Arg	Pro	Pro	(13)
E. coli[4] HLT A Subunit	14	(6)	Tyr	Arg	Ala	Asp	Ser	Arg	Pro	Pro	(13)
Region 2											
Pertussis S1 subunit	15	(51)	Val	Ser	Thr	Ser	Ser	Ser	Arg	Arg	(58)
Cholera[3] A subunit	16	(60)	Val	Ser	Thr	Ser	Ile	Ser	Leu	Arg	(67)
E. coli[4] HLT A Subunit	17	(60)	Val	Ser	Thr	Ser	Leu	Ser	Leu	Arg	(67)

The numbers in parentheses refer to the amino acid position in the mature proteins.
[1]Data from Yamamoto, et al. FEBS Letter 169:241, 1983
HLT – Heat Labile Toxin --

Please amend and replace the paragraph, at page 26, line 22 through page 27, line 25, as follows:

-- Since all pertussis toxin subunits are closely linked and probably expressed in a very precise ratio, it is possible that they are arranged in a polycistronic operon. A polycistronic arrangement for the subunit cistrons also has been described for other bacterial toxins bearing similar enzymatic functions, such as diphtheria, cholera, and E. coli heat labile toxins. Therefore, the flanking regions were analyzed for the presence of transcriptional signals. In the 5' flanking region, starting at position 469, the sequence TAAAATA (SEQ ID NO:18) was found, which six of the seven nucleotides found in the ideal TATAATA (SEQ ID NO:19) Pribnow or –10 box. An identical sequence can be found in several other bacterial promoters, including the lambda L57 promoter. Given the fact that most transcripts start as a purine residue about 5-7 nucleotides downstream from the Pribnow box, the transcriptional start site was

tentatively located at the adenine residue at position 482. This residue is located in the sequence CAT, often found at transcriptional start sites. Upstream from the proposed –10 box, the sequence CTGACC (SEQ ID NO:20) starts at position 442. This sequence matches four of the six nucleotides found in the ideal E. coli -35 box TTGACA (SEQ ID NO:21). The mismatching nucleotides in the proposed pertussis toxin –35 box are the two end nucleotides, of which the 3' residue is the less important nucleotide in the E. coli –35 consensus box. A replacement of the T by a C in the first position of the consensus sequence can also be found in several E. coli promoters. The distance between the two proposed promoter boxes is 21 nucleotides, a distance of the same length has been found in the galP1 promoter and in several plasmid promoters. The proposed –35 box is immediately preceded by two overlapping short inverted repeats with calculated free energies of –15.6 kcal and –8.6 kcal, respectively. Inverted repeats can also be found at the 5'-end of the cholera toxin [promotor] promoter. In both cases, they may be involved in positive regulation of the toxin promoters. None of the ORFs assigned to the other subunit is closely preceded by a similar promoter-like structure. However, a different promoter-like-structure was found associated with the S4 subunit ORF. --

Please amend and replace the paragraph, at page 28, line 15 through page 29, line 12, as follows:

-- Additionally, the 5'-flanking region of each cistron was also examined for the presence of ribosomal binding sites. Neither the ribosomal binding sequences for B. pertussis genes, nor the 3'-end sequence of the 16S rRNA are known. Therefore, the flanking regions could be compared with only the ribosomal binding sequences of heterologous procaryotic organisms represented by the Shine-Dalgarno sequence. Preceding the S1 initiation codon, the sequence GGGGAAG (SEQ ID NO:22) was found starting at position 495. This sequence shares four out of seven nucleotides with ideal Shine-Dalgarno sequence AAGGAGG (SEQ ID NO:23). The two first mismatching nucleotides in the pertussis toxin gene would not destabilize the hybridization to the 3'-end of the E. coli 16 S rRNA. This putative ribosomal binding site is close enough to the initiation codon for S1 to be functional in E. coli. Another possible Shine-Dalgarno sequence overlaps the first one and also matches four out of seven nucleotides to the

consensus sequence. The mismatching nucleotides, however, have a more destabilizing effect than the ones found in the first sequence. The S2 subunit ORF is not closely preceded by a ribosomal binding sequence, which may suggest that S2 is translated through a mechanism not involving the detachment and reattachment of the ribosome between the coding regions for S1 and S2. The short distance between the S1 and S2 cistrons, and the absence of a ribosomal binding site are characteristic of this mechanism. A ribosomal binding site for S4 in the sequence CAGGGCGGC (SEQ ID NO:24), starting at position 2066 is possible. The ORF for S5 is preceded by the sequence AAGGCG (SEQ ID NO:25), starting at position 2485, which matches five out of six nucleotides in the consensus sequence AAGGAG (SEQ ID NO:26). Finally, S3 is preceded by the sequence GGGAACAC (SEQ ID NO:27), which is very similar to the proposed ribosomal binding site for S1, i.e., GGGAAGAC (SEQ ID NO:28).--